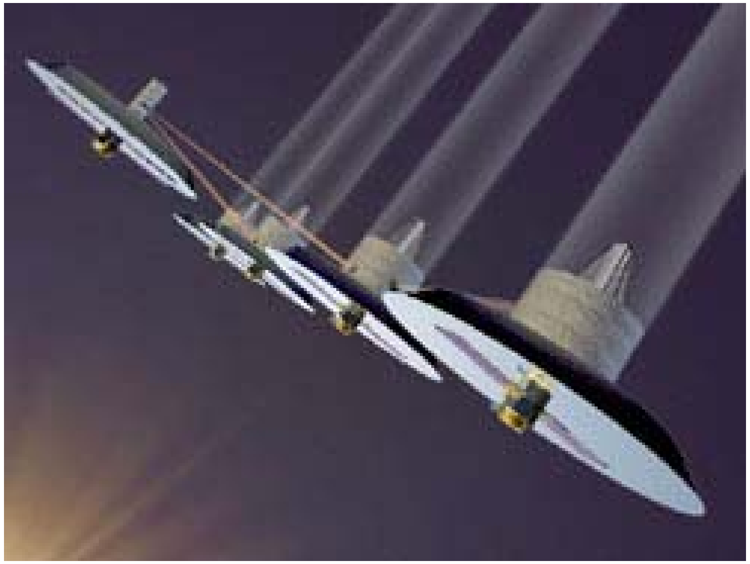
Science Models: How Uncertainty Drives Design



Dr. Raymond Sedwick
Principal Research Scientist

Dan Kwon, SM Candidate
Graduate Research Assistant

Presented by Julie Wertz
PhD Candidate
Michelson Research Fellow

Massachusetts Institute of Technology
Space Systems Laboratory

Bounding $\eta_\oplus$ (What is η_u?)

- At a previous SWG, some discussion of how $\eta_\oplus$ impacts mission success
- Statement made that if value of $\eta_\oplus$ is too small, there is no hope of success due to finite mission life
  - But, what if we know exactly which stars to look at? (Precursor science)
- There are two separate issues:
  1. How many stars must we look at if we hope to find 'Earths'
  2. How far must we be able to look to see the 'right' stars
- We want to address the question of how precursor science, as well as uncertainty, impacts system design

- Definitions used:
  - $\eta_\oplus$ – Fraction of TOTAL stars that HAVE planets of 'interest' (POI)
  - η_u – Fraction of TOTAL stars we THINK have planets of interest. This is the upper bound for $\eta_\oplus$
  - p – Probability of finding POI within a sample of stars
  - N_S – Total number of stars in a volume of space
  - R – Radius of volume of space
  - $N_{S/P}$ – Total number of stars with POI within a volume of space

Why is η_u an important parameter?

- By its definition, η_u must vary between the value of $\eta_\oplus$ and 1

$$\eta_\oplus \le \eta_u \le 1$$

- If we assume that we are not going to sample stars that are unfavorable, then the probability (*p*) of finding a POI within a sample of stars is

$$p = \frac{\eta_\oplus}{\eta_u}$$

- With no a priori knowledge (precursor), this probability is simply $\eta_\oplus$, but with complete knowledge it could equal 1

- If only 1 in 1000 stars has a POI, but we know which one it is, then we only have to look at one star
  - BUT it still may be very far away

- Although η_u factors into how many stars we need to search, it does not affect how far we must look
- Assuming stars are uniformly distributed:

$$N_{S/\oplus} = \eta_\oplus N_S = \eta_\oplus \frac{4}{3}\pi R^3 n_S$$

$$R = \left(\frac{3N_{S/\oplus}}{\pi n_S \eta_\oplus} \right)^{1/3}$$

For example, if $\eta_\oplus$ = 0.001
and η_u = 0.01 (good knowledge)

We have a 0.1 probability that the stars we look at will have planets, so we don't have to survey as many of them

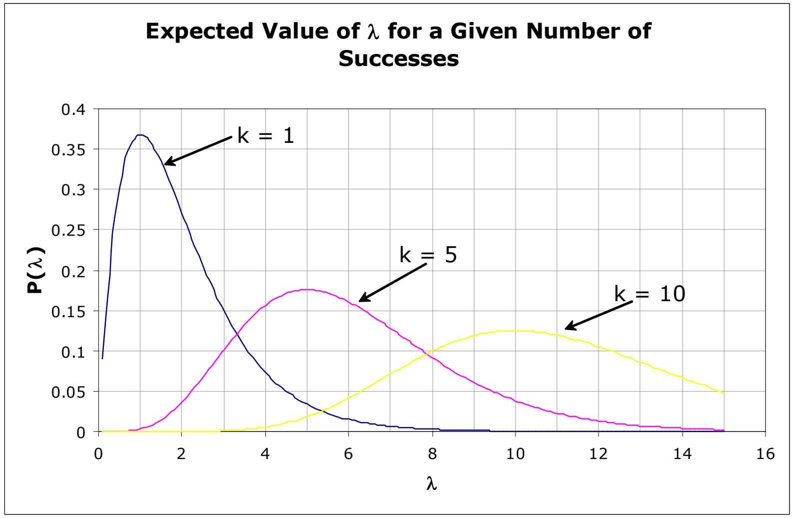
BUT we have to look 10x farther to find them

- Impact: May still need larger baselines

Applying Sampling Statistics

- For large sample size (N), the sampling statistics approach a Poisson process

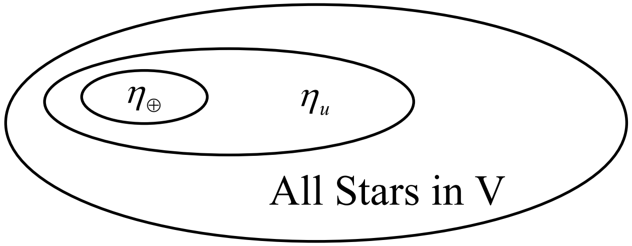
$$P(k)=\frac{\lambda^k}{k!}e^{-\lambda} \qquad \lambda = pN$$



- The expected value of λ for the distribution is the number of trial successes
- Given *N* samples, and *k* successes, the most likely value of *p* is *k/N*
- The uncertainty in λ is smaller with fewer successes

- Only F, G and K type stars are assumed to be likely to have planets of interest
- Currently our 'guess' at η_u is then

$$\eta_u = \frac{N_{F-G-K}}{N_{All}}$$



$\Rightarrow$ We are only considering those stars

- If precursor science 'surveys' 100 F-G-K stars, and 'identifies' 2 POIs, then the most likely value of λ for the sample is 2, and p=1/50

- But, what if we also learn something about why these stars had planets?

How Can Precusors Impact η_u?

- If precursor observation and theory estimates that 1/50 of F-G-K type stars has POI, then two things might happen:

1) If a notable correlation can be made, then η_u can be reduced, and *p* increased
   - E.g. 'All favorable stars are G-type'

2) Otherwise, we don't know how to be more selective in our search, and therefore cannot reduce η_u
   - Instead, we know that the p associated with our current search strategy has an expected value of

$$p = \frac{\eta_\oplus}{\eta_u} \leq 0.02$$

- If P_f is the likelihood of failing to find any planets in the next sample set, we can calculate

$$N_{samples} = -\frac{1}{p}\ln(P_f)$$

- Given *p*, we can determine the number of additional samples that must be taken
- With the current value of p = 0.02, a likelihood of 1% for failing to find any planets of interest would mean we have to survey >230 stars
- If it is eventually determined that there is a reason why only these 1/50 stars has a POI, then the value of η_u can be reduced by a factor of 1/50, as our future search strategies will be more selective

When Does η_u Change?

- η_u changes as a result of processed precursor science mission data
- Data processing will lag launch dates by 2-3 years

2004 2005 2006 2007 2008 2009 2010 2011 2012 2013

SIRTF
SOFIA

Select Baseline Architecture

LBTI

Kepler
-indirect planet detection using "transition" method

Eclipse

SIM

JWST

→TPF

- Current Projects
- Keck (July 2003 – 1st published science observation, 2001 – 1st light)
  – Exozodiacal dust characterization & nulling
  – Can detect Uranus size planets around stars up to 60 light yrs away

- Future Projects before Architecture Selection
  – SOFIA
    - Proto-planetary disks, planet formation in nearby star systems
  – SIRTF
    - Brown dwarfs, super-planets
    - Dust disks surrounding nearby stars

Most of the Precursor Science Data Arrives AFTER Baseline Architecture Selection

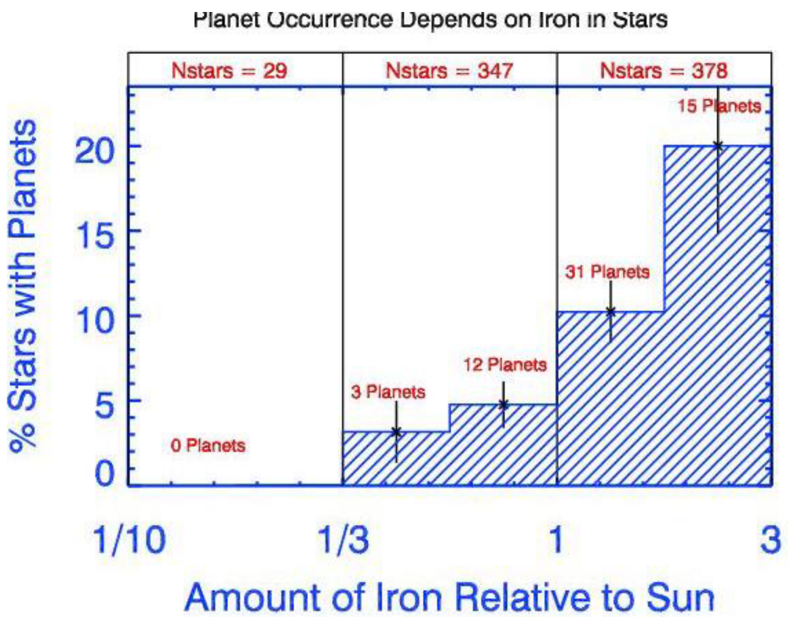
How Does η_u Change? (Case Study)

- η_u changes with significant precursor science/observation
- **Example**: Correlation between iron-rich stars and large planets*
- **Scenario** – 754 nearby sun-like stars surveyed to determine the presence of a Jupiter-sized planet & spectrum of each star taken to determine amount of iron
- **Results** - 61 planets existed. Data suggests metal-rich stars are more likely to develop planetary systems
  - Near linear relationship between Fe and planet existing

- How this affects η_u for large planets around sun-like stars:
  - Stars with 1/10 to 1/3 the amount of Fe appear to have no detectable planets
  - Assuming this data is representative of an arbitrarily large data set

$$\eta_u^{new} = \eta_u^{old}\left(1 - \frac{29}{754}\right)$$

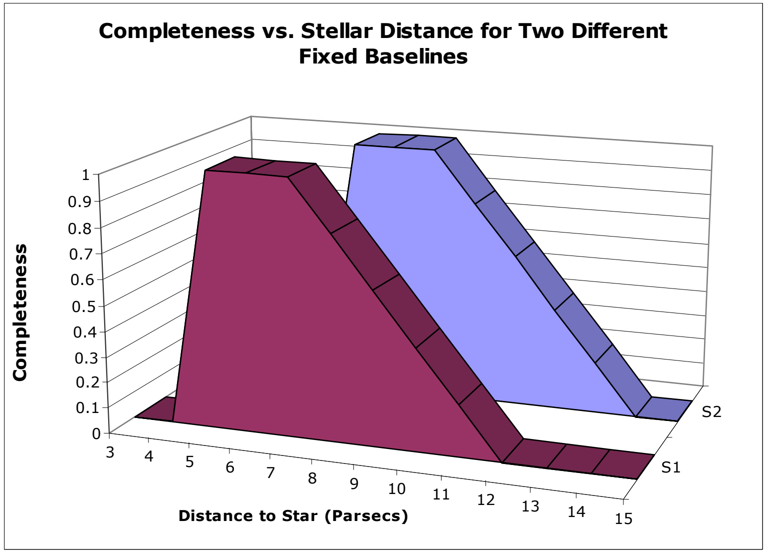
  - Incorporating the weighted probabilities is work in progress



*Fischer, UCBerkeley/Valenti, STScI, http://exoplanets.org/metalicity.html

Limits of a Fixed Baseline System

- The time required to do detection about a given star is the same regardless of habitable zone coverage (one aspect of ‘completeness’)
  - Less than 100% coverage allows possibility of missing a planet that is there
- Fixed baseline systems will have 100% coverage over some stars, but partial coverage over most.
  - Stars with 100% coverage occur in a narrow distance band
  - A larger number of stars can be included in this band by looking farther out, but at the expense of longer integration times (fewer stars actually viewed)



	BL 1	BL 2	BL 3
# of stars in ΔR (C=1)	100	500	1000
# viewable over mission	50	40	30
# of planets if $\eta_\oplus = 0.01$	1	5	10
P_f if $\eta_\oplus = 0.01$, $\eta_u = 1$	60%	67%	74%
# of planets if $\eta_\oplus = 0.001$	0	0	1
P_f if $\eta_\oplus = 0.001$, $\eta_u = .01$	100%	100%	5%

- Higher uncertainty in which stars should be observed favors looking closer, since a larger sample set can be viewed over the mission lifetime
  - This is the effect of distance on integration time
- However, if $\eta_\oplus$ is much smaller than predicted, the number of stars accessible by a shorter fixed baseline system may be insufficient to have an ‘Earth’
- Longer fixed baseline systems are less likely to find Earths if η_u is large, but may be the only chance of finding them if precursor missions reduce η_u

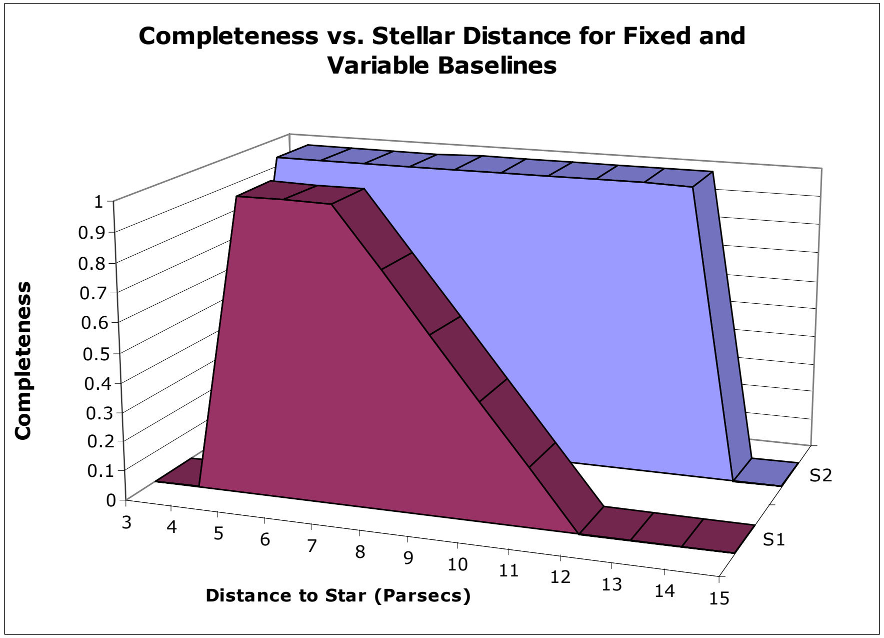
Because precursor science data will lag the architecture downselect there is no way to know whether a shorter or longer fixed baseline is better

Merits of a Variable Baseline

- Unlike a fixed baseline system, a variable baseline allows for observational completeness regardless of the stellar distance

- If η_u turns out to be relatively large, operating at shorter baselines will allow for many more observations to be made over the mission lifetime, increasing science throughput

- If η_u turns out to be relatively small, operating at longer baselines grants access to a much larger volume of stars, increasing science throughput



- If η_u turns out to be very small, then very large baselines may be necessary to access a sufficient volume of space to insure the stars with planets of interest can be observed
  - The longer integration times make the need for a tunable baseline (to achieve full completeness) even more important

Independent of technological issues, for or against, a formation flown variable baseline system appears to offer the lowest risk for science return

Summary

- Even if η_u is very small, precursor science will allow us to upper bound its value in such a way that we can be smarter about which stars are observed
- A large uncertainty in η_u drives us toward a desire for more observations, and therefore shorter baselines
  - The danger lies in the possibility that the limited number of stars accessible by a shorter fixed baseline system may no have any ‘Earths’
- Larger fixed baseline systems allow for more stars to be accessed, but without good precursor knowledge, we cannot effectively limit our search
- The problem is that the architecture downselect occurs before it will be known whether searching closer or farther is the better approach
- The most robust solution is to make the baseline variable, over a relatively large extent, allowing operation in either mode
  - This supports a formation flown system as providing the least science risk